TWO LAYER VECTOR QUANTIZATION OF IMAGES

Ján MIHALÍK, Iveta GLADIŠOVÁ, Viktor MICHALČIN
Dept. of Electronics and Multimedia Communications
Technical University of Košice
Park Komenského 13, 041 20 Košice
Slovak Republic

Abstract

This paper deals with spatial scaleable vector quantization of images by using two layers. At the base layer the small dimensional full search vector quantization of an image of block means is carried out. After subtracting it from the input image at the enhancement layer the difference of full spatial resolution is quantized by large dimensional weighted pyramid vector quantizer in transform domain. Experimental results demonstrate acceptable quality of the downscaled image from the output of base layer at very low bit rate and possibility of next increasing its quality by the upscaled difference image from the output of enhancement layer.

Keywords

Base layer, enhancement layer, full search vector quantization, weighted pyramid vector quantization

1. Introduction

Vector quantization (VQ) [1], [2] has become a popular and powerful data compression technique for image coding. Although the performance of full search VQ (FSVQ) [9] approaches the theoretical upper bound as the vector dimension increases, the complexity involved has made it impossible to use large dimensionality FSVQ in practice. Due to the unstructured nature of the codebook generated by LBG algorithm [4] the encoding complexity generally increases exponentially with rate (for a fixed dimension) or dimension (for a fixed rate). Several sub-optimum VQ techniques [5], [6] have been proposed which mitigate the complexity barrier. In this paper we propose two layers VQ (2LVQ) of images, that enables vector quantization of large image blocks with tolerable encoding complexity and permits progressive image reconstruction. The 2LVQ uses the concept of small dimensional FSVQ at the base layer and large dimensional weight pyramid VQ (WPVQ) [7] in transform domain at the enhancement layer.

2. Basic idea of 2LVQ

Block scheme of 2LVQ is in Fig. 1. After segmenting of the input image to the blocks of size 4x4 pixels, sequence of the blocks (vectors) is fed into 2LVQ. By using average operation the mean value is calculated for each block. Next four means of the neighbouring blocks ordered inside of the macro-block of size 8x8 pixels are grouped to one four dimensional vector. The sequence of mean vectors is vector quantized at the base layer by the full search VQ. From its output we get quantized block means, which create an image with four times lower spatial resolutions in both directions compared to those of the input image. At the enhancement layer the image of the block means is subtracted from the input image by the way that same mean value is subtracted from all pixels inside of its block. From the output of subtractor we get partially decorrelated difference image with full spatial resolution. Afterwards difference image is segmented to the sequence of difference blocks of the size 4x4 pixels, which are transformed using DCT. Statistical properties of the blocks in transform domain are better approximated by multiple Laplacian probability distribution with zero means and nonlinear decomposed variances of transform coefficients. Therefore transformed difference blocks are vector quantized in sixteen dimensional weighted pyramid VQ that is very good adapted to the probability low. Reconstruction of the difference image from the output of enhancement layer is carried out by using inverse DCT. If we add reconstructed difference image to the image of block means from the base layer interpolated by holding the means over their blocks we get the complete reconstructed image of the high quality with full spatial resolution.

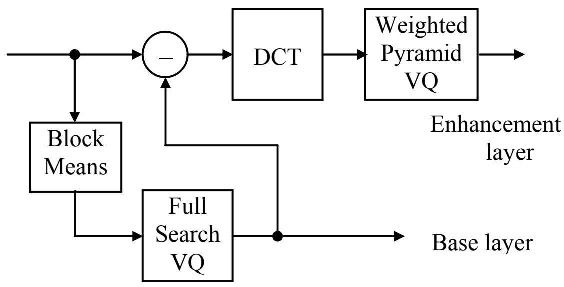


Fig. 1 Block scheme 2LVQ

3. Full search vector quantization

The basic quantization problem considered in the sequel is the following. Let **X** be an v-dimensional random vector, **Q** denotes a mapping of a vector quantizer and **Y** = **Q(X)** its output with $\mathbf{b}_i$, i = 1, ..., N, the N possible quantizer output vectors. The mapping **Q** is to be selected to minimize the mean square error per dimension [8].

$$\sigma_q^2 = \frac{1}{v} E\left(\left\|\mathbf{X} - \mathbf{Q}(\mathbf{X})\right\|^2\right), \tag{1}$$

where E is statistical mean value.

Necessary conditions for an optimum FSVQ may be easily derived. In summary, associated with each quantizer output vector $\mathbf{b}_i$ is a nearest neighbour partition cell A_i, i = 1,..., N, satisfying

$$A_i = \left\{\mathbf{X} : \left\|\mathbf{X} - \mathbf{b}_i\right\|^2 < \left\|\mathbf{X} - \mathbf{b}_j\right\|^2, \quad \forall\, j \neq i\ \right\} \tag{2}$$

For each partition cell A_i, $\mathbf{b}_i$ is the conditional mean

$$\mathbf{b}_i = \mathrm{E}\left(\mathbf{X} / \mathbf{X} \in \mathrm{A}_i\right) \tag{3}$$

General sufficiency conditions for optimality in the vector case are not known, and there may exist many locally optimum VQs that satisfy (2) and (3).

The necessary conditions (2) and (3) provide bases of iterative LBG quantizer design algorithm. Since analytical or numerical computation of (3) is not feasible, a random number generator is used to produce a training sequence ($\mathbf{x}_n$, n = 1,... , T), representing the distribution of random vector **X** (or, for a physical source of interest such as image, an appropriate training sequence is selected from available data). Given an initial set of output vectors ($\mathbf{b}_i$, i = 1, ..., N), (2) can be applied to the training sequence and various $\mathbf{x}_n$ assigned to cells A_i of partition. Eqn. (3) can then be solved for a new set of quantizer output vectors. These output vectors are in general different from the initial values and can be used again in (2) with the training sequence to reassign each $\mathbf{x}_n$ to a new partition. Iterating between (2) and (3) provides a non-increasing distortion, and the algorithm eventually converges to a locally optimum design.

The basic idea of the LBG algorithm [4] is quite simple and can be summarized for an unknown distribution training sequence as follow:

1. Let N is number of quantizer output vectors, distortion threshold $\varepsilon \geq 0$, assume an initial N-vector reproduction alphabet B_0, and training sequence $\mathbf{x}_n$, n = 1, ..., T, and r is the number of iterations set to zero.
2. Given $B_r = (\mathbf{b}_i$, i = 1, ..., N), find the minimum distortion partition $P(B_r) = (A_i$, 1 = 1, ..., N) of the training sequence : $\mathbf{x}_n \in A_i$ if $d(\mathbf{x}_n , \mathbf{b}_i) \leq d(\mathbf{x}_n , \mathbf{b}_j)$, for all j, where d is distance. Compute time average distortion

$$\sigma_{qr}^2 = \sigma_q^2\left[\mathrm{B}_r, \mathrm{P}(\mathrm{B}_r)\right] = \frac{1}{\mathrm{T}} \sum_{n=1}^{\mathrm{T}} \min_{b_i \in \mathrm{B}_r} d(\mathbf{x}_n, \mathbf{b}_i)\,. \tag{4}$$

3. If $(\sigma^2_{qr-1} - \sigma^2_{qr}) / \sigma^2_{qr} \leq \varepsilon$, stop the iteration with the B_r final reproduction alphabet, otherwise continue.
4. Find the optimum reproduction alphabet $B[P(B_r)] = [\ \mathbf{b}_i(A_i)$, i = 1, ..., N] for $P(B_r)$ where

$$\mathbf{b}_i(\mathrm{A}_i) = \frac{1}{\left\|\mathrm{A}_i\right\|} \sum_{n:\mathbf{x}_n \in \mathrm{A}_i} \mathbf{x}_n \tag{5}$$

5. Set B_{r+1} = B [P (B_r)] and increment r to r+1 and go to item 2.

In the above iterative algorithm, an initial reproduction alphabet B_0 was assumed in order to start the algorithm. There are a number of techniques to obtain the initial reproduction alphabet (codebook). The simplest technique is to use the first widely spaced words from the training sequence. The another technique uses a splitting method where the centroid for the training sequence was calculated and split into two close vectors. The centroids or the reproduction vectors for the two partitions were then calculated. Each resulting vector was then split into two vectors and the above procedure was repeated until an N-vector initial reproduction alphabet was created. Splitting was performed by adding a fixed perturbation vector **k** to each vector $\mathbf{b}_i$ producing two vectors $\mathbf{b}_i + \mathbf{k}$, $\mathbf{b}_i - \mathbf{k}$.

4. Weighted pyramid vector quantization

In the source coding context, from the rate-distortion theory it follows that almost all code-words (reproduction vectors) are selected to lie in a region of high probability [9] specified by the entropy of the source. This region of high probability will have a geometry that is dependent on the source (e.g., a sphere for memoryless Gaussian source, a hypercube for a uniform source, a pyramid for a memoryless Laplacian source). Thus good codebooks can also be constructed as the intersection of the points in the lattice and the geometric region of high probability for source.

The vector quantizers designed in such a way have a simple algorithm of code-words searching. The weighted pyramid vector quantizer is suitable for the sequence of vectors $\mathbf{X}' = (X_1', X_2', \ldots, X_v',)^T$ with Laplacian probability distribution while their components X_i' are statistically independent and have different variances σ^2_{xi} or parameter $\lambda_i = 2^{1/2} / \sigma_{xi}$ for i = 1, ..., v and let $\lambda^T = (\lambda_1, \lambda_2, \ldots, \lambda_v)$. The probability density function for such a random vector is

$$f(\mathbf{x}) = \prod_{i=1}^{v} \frac{\lambda_i}{2} e^{-\lambda_i \left|x_i'\right|}, \tag{6}$$

then the weighted pyramid is defined as follows

$$S_{WP}(v, \lambda) = \left\{\mathbf{x} : \frac{1}{v} \sum_{i=1}^{v} \lambda_i \left|x_i'\right| = 1\right\} \quad . \tag{7}$$

The vector quantization, which searches the code-words from the pyramid, determined by equation (7) is a weighted pyramid vector quantization, which means that the pyramid $S_{WP}(v, \lambda)$ is the optimum one for the WPVQ design.

On the basis of the comparison of relationships (6) and probability density function for non-weighted pyramid vector quantizer (NWPVQ) [10], the equation $\lambda x_i = \lambda_i x_i'$ can be obtained. Then for components x_i' of the input vector a mapping relationship is valid

$$x_i' = \frac{\lambda}{\lambda_i} x_i = \frac{1}{w_i} x_i \quad , \tag{8}$$

where $w_i = \lambda_i / \lambda$ is the weight of appropriate components of the vector **X'**. The relationship allows transforming the WPVQ to NWPVQ substituting (8) into (7) as follows

$$\frac{1}{v}\sum_{i=1}^{v} \lambda_i \left| \frac{\lambda}{\lambda_i} x_i \right| = \frac{1}{v}\sum_{i=1}^{v} \lambda |x_i| = 1 \quad , \tag{9}$$

and from this $\sum_{i=1}^{v} |x_i| = v/\lambda$, which corresponds to the equation of the optimum pyramid of the NWPVQ while

$$\lambda = \left(\prod_{i=1}^{v} \lambda_i \right)^{1/v} \quad . \tag{10}$$

WPVQ distortion [5] for a large dimension v and bit rate n is

$$\begin{aligned} \sigma_{q\,WP}^2(n) &\cong \frac{e^2}{3}\left[\prod_{i=1}^{v} \frac{1}{\lambda_i^2} \right]^{1/v} 2^{-2n} = \\ &= \frac{e^2}{3}\left[\prod_{i=1}^{v} \frac{\sigma_{xi}^2}{2} \right]^{1/v} 2^{-2n} \quad . \end{aligned} \tag{11}$$

It is required to point out that the WPVQ to NWPVQ transformation also allows to utilize the NWPVQ algorithm at the design of the WPVQ algorithm [7].

Fig. 2a shows the non-weighted pyramid S (3,4) and Fig. 2b shows the weighted pyramid S_{WP} (3,4) for the weight vector $\mathbf{w} = (2^{1/2}/2;\ 1;\ 2^{1/2})$ with marked code-words at upper portion of the pyramid.

For our algorithm of the weighted pyramid vector quantizer we will utilize the possibility of the WPVQ to the NWPVQ mapping and hence the transform of the NWPVQ algorithm to the WPVQ algorithm, which contains

1. Transformation of the input vector **X'** to the vector **X**.
2. Simulation of the algorithm of the NWPVQ [11], i.e. searching for the codeword to the vector **X**.
3. Inverse transformation of the selected codeword to the weighted pyramid and hence obtaining the resulting codeword of the WPVQ.

This results in the fact that the number of WPVQ code-words, designated as N(v,K), where $v/\lambda = K$ and K is a positive integer number, will be equal to the number of NWPVQ code-words and it is given by relationship [10]

$$N(v, K) = N(v-1, K) + N(v-1, K-1) + N(v, K-1) \quad \text{for } v \ge 2, \quad K \ge 2 , \tag{12}$$

with the boundary conditions N(v, 1)=2v and N(1, K) = 2.

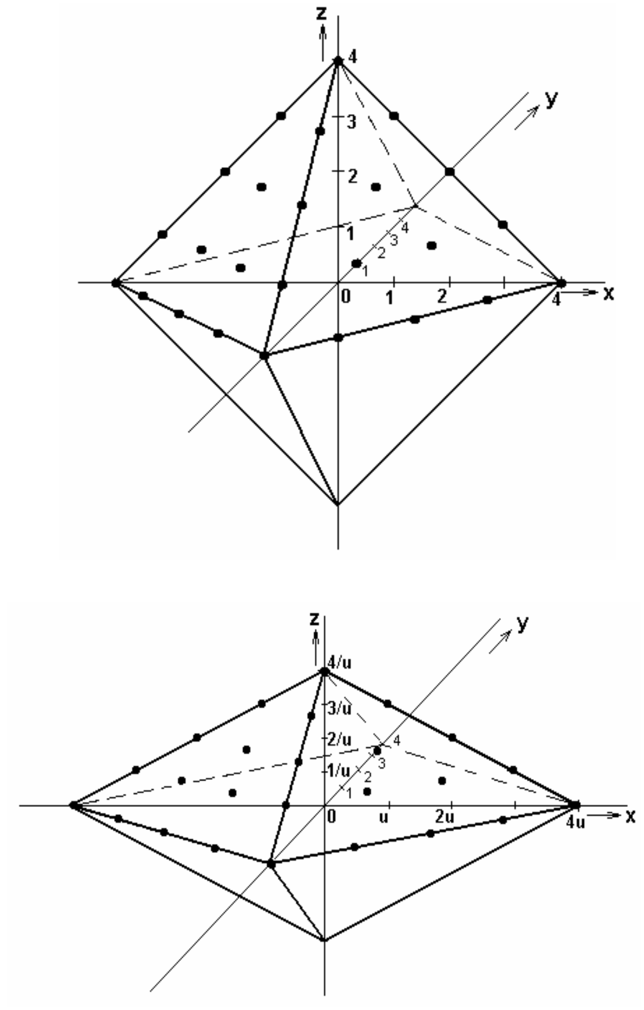


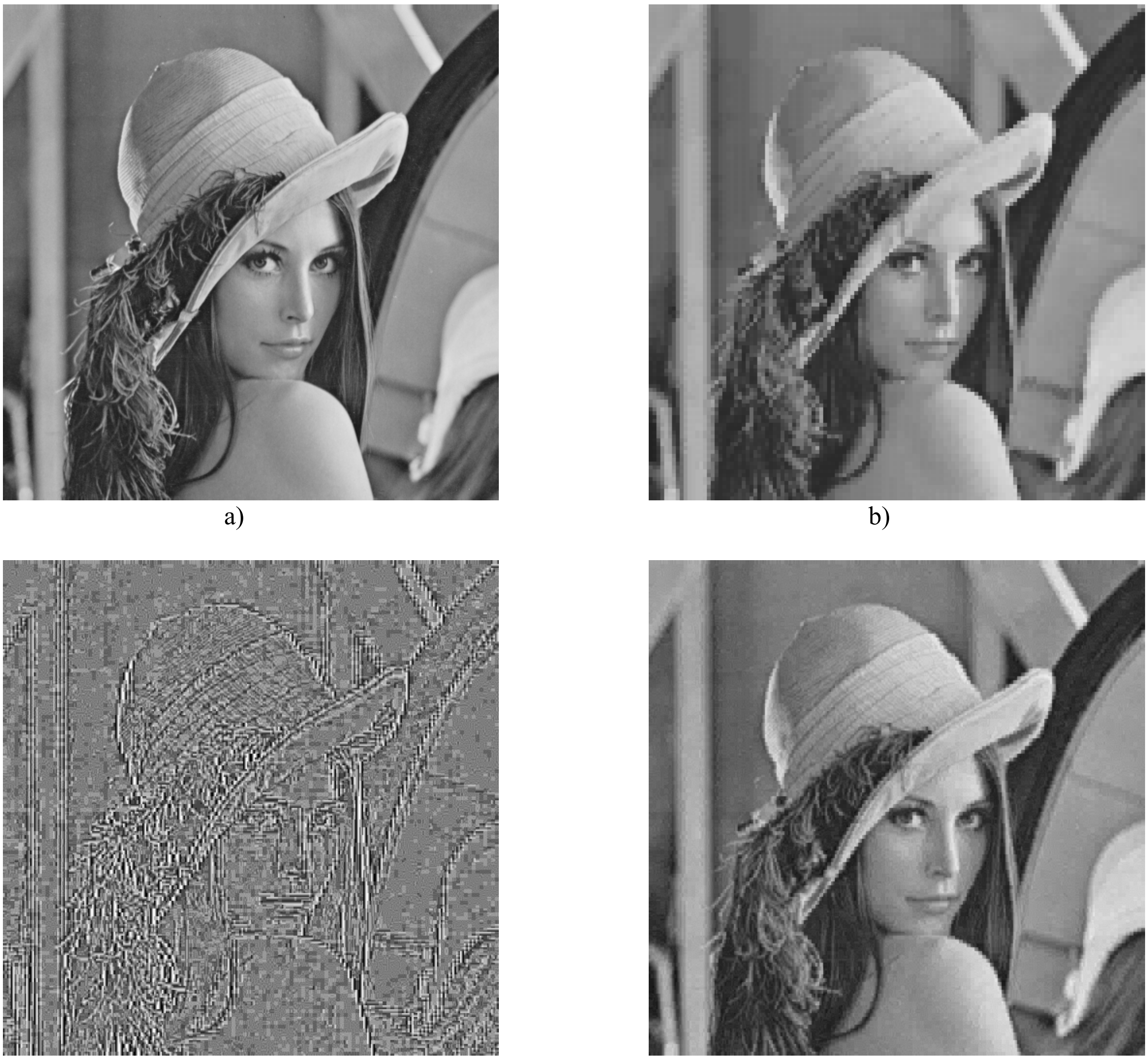
Fig. 2 The relationship of non-weighted and weighted pyramids in the three-dimensional space and distribution of code-words at a) non- weighted pyramid; b) weighted pyramid, where $u = \sqrt{2}$.

The algorithm of searching for the WPVQ code-words at a bit rate n bits/dimension requires the specification of the largest K such that $N(v,K) \le 2^{nv}$.

The basic encoding algorithm requires the addition, multiplication, comparison and rounding operations and its computational complexity is proportional to dimension v, while for FSVQ it increases exponentially with dimension.

5. Simulation results

Computer simulations were carried out using the proposed 2LVQ and design algorithms of FSVQ and WPVQ to scaleable encode monochrome images of 512x512 pixels with 256 grey levels in Fig. 3a. At the base layer the input image is encoded in decimated form of size 128x128 pixels at very low bit rate 0,125 bits/pel. In regard strong correlation of the decimated image represented by block means it was vector quantized by 4-dimensional 8 bit FSVQ. The output image of acceptable quality from the base layer is in Fig. 3b and is displayed on the image plane of same size as that of the other decoded full spatial resolution images. At the enhancement layer the difference image between input and vector quantized image from the base layer was encoded with full spatial resolution. The difference image is partially decorrelated but its spatial correlation is still enough large to directly apply on it pyramid VQ.



a) b) c) d)

Fig. 3 a) Original image of Lenna of the size 512x512 pixels. Decoded images from b) base layer of the size 128x128 pixels at 0,125 bits/pel. c) enhancement layer of the size 512x512 pixels at 0,5625 bits/pixel. d) final decoded image of the size 512x512 pixels at 0,6875 bits/pixel.

Therefore to better approximate its statistical properties by multiple Laplacian probability distribution of independent components with zero mean values and nonlinear decomposition of their variances we first transform it by 4x4 DCT and afterwards WPVQ is carried out. The transform difference image is vector quantized by 16-dimensional WPVQ at bit rate 0,5625 bits/pixel and after the inverse DCT is shown in Fig. 3c. By summing the both images we get the final decoded full spatial resolution image in Fig. 3d with high quality at the total bit rate 0,6875 bits/pixel. As we can see the final decoded image contains except for block means high frequency components. In addition block means have smaller distortion compared to those of the base layer because they are finer quantized at the enhancement layer, too.

6. Conclusion

In this paper, a combination of the full search and weighted pyramid VQ has been presented at the two-layer structure. The structure consists the small dimensional FSVQ at the base layer and the large dimensional WPVQ in DCT domain at the enhancement one. With regard to small dimension of FSVQ and the geometrical properties of WPVQ, the designed structure is very simple and fast suitable for real time vector quantization of images. Experimental results of spatial scalelable vector quantization of images by using the two layer structure demonstrate acceptable quality of downscaled image from the output of base layer at very low bit rate and possibility of next increasing its quality by the upscaled difference image from the output of enhancement layer. By applying of entropy encoding above all at the base layer we can achieve higher performance of the 2LVQ.

Acknowledgements

The presented work was supported by the Grant Agency of Ministry of Education and Academy of Science of Slovak Republic under Grant No. 1/7387/20.

References

[1] NASRABADI, N.M., KING, R.A. Image Coding Using Vector Quantization: A Review. IEEE Transactions on Communications, vol. 36, no. 8 ,1988, pp. 957-971.

[2] MIHALÍK, J. Hierarchical Vector Quantization of Images in Transform Domain. Electrical Engineering Journal, vol. 43, no. 3, 1992, pp. 92-94.

[3] CHMÚRNY, J., MIHALÍK, J. Algorithms for Optimisation of Vector Quantizers. Slaboproudý obzor, vol. 47, no. 9, 1986, pp. 435-440. (In Slovak)

[4] LINDE, Y., BUZO, A., GRAY, R.M. An algorithm for Vector Quantizer Design. IEEE Transactions on Communication, vol. 28, no. 1, 1980, pp. 84-95.

[5] GERSHO, A., GRAY, R.M. Vector Quantization and Signal Compression. Kluwer Academic Publishers, Dordrecht, Notherlands, 1992.

[6] MIHALÍK, J. Multistage Vector Quantization of Image by Using Clustering Interpolation in DCT Domain. Radioengineering, vol. 6, no. 4, 1997, pp. 10-13.

[7] MIHALÍK, J., GLADIŠOVÁ, I. Weighted Pyramid Vector Quantizer. Electrical Engineering Journal, vol. 46, no. 2, 1995, pp. 46-50.

[8] MIHALÍK, J. Digital Signal Processing. Alfa Bratislava, 1987. (In Slovak)

[9] MIHALÍK, J. Contour Based Scalar-Vector Quantizer. Electrical Engineering Journal, vol. 46, no. 4, 1995, pp. 121-125.

[10] FISCHER, T.R. A pyramid Vector Quantizers. IEEE Transactions on Information Theory, vol. 32, no. 4, 1986, pp. 568-583.

[11] GLADIŠOVÁ, I., MIHALÍK, J. An Algorithm of Pyramid Vector Quantization. In Proceedings of the International Conference "Digital Signal Processing", Košice, 1993, pp. 236-239. (In Slovak)

About authors...

Ján MIHALÍK was born on August 1, 1952 in Slovakia. He graduated from Technical University in Bratislava in 1976. Since 1979 he joined Technical University of Košice, where received his PhD degree in Radioelectronics in 1985. Currently, he is Full Professor at the Department of Electronics and Multimedia Telecomunications of Technical University, Košice. His research interests include information theory, digital image processing and video-communication.

Iveta GLADIŠOVÁ was born in Slovakia. She graduated from Technical University of Košice in 1984. Since 1984 she joined same university of Košice, where received his PhD Degree in Radioelectronics in 1997. She is Assistant Professor at the Department of Electronics and Multimedia Telecomunications of Technical University, Košice. Her research interests include digital image processing and vector quantization.

Viktor MICHALČIN was born on September 6, 1976 in Ukraine. He received the Ing. Degree from Technical University of Košice in 2000. Currently he is PhD student at the Department of Electronics and Multimedia Telecomunications of Technical University, Košice. His research interests are on vector quantization of signals.